УДК 519.2

ИДЕНТИФИКАЦИЯ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗМЕРЕНИЙ ЭКОНОМИЧЕСКИХ ПАРАМЕТРОВ НА ОСНОВЕ СКРЫТОЙ МАРКОВСКОЙ МОДЕЛИ

С.В. Шалагин,
Казанский национальный исследовательский технический университет
им. А.Н. Туполева – КАИ, г. Казань
А.Р. Нурутдинова,
Набережночелнинский институт (филиал) Казанского (Приволжского)
федерального университета, г. Казань

Ключевые слова: *алгоритм идентификации, скрытые марковские модели, циклические цепи Маркова, автоматная марковская модель.*

В теоретических и прикладных исследованиях в различных областях науки и техники [1–4], в частности, в области анализа состояния экономических объектов и систем важную роль играет задача идентификации автоматных марковских моделей (АММ). Задача основана на сборе последовательных измерений выходных параметров объекта, при этом наблюдаемая последовательность рассматривается как случайный процесс, описывающий изменения состояний системы в дискретные моменты времени, и расчете вероятности отнесения объекта к тому или иному априори заданному классу.

Один из методов идентификации был предложен Л.Р. Рабинером в рамках решения задачи оценивания вероятности некоторой последовательности наблюдений, определяемой заданной скрытой марковской моделью, наилучшим образом соответствующей наблюдаемому измерению [5]. В работе предложена модификация данного алгоритма, которая применима к решению задачи идентификации автоматной марковской модели, заданной на базе простых стохастических матриц класса эргодических регулярных (ЭСМ) и циклических ($ЦСМ_r$). Идентификация производится на основе последовательности измерений экономических параметров заданной длины *N*. Причем часть элементов анализируемой последовательности могут быть не определены.

Пусть простая однородная цепь Маркова (ЦМ) задана в виде [6]

$$(S, P_{(m)}, \pi_0), \tag{1}$$

где $S = \{s_i\}$, $i = \overline{0, n-1}$ – множество состояний ЦМ, P_S – ЭСМ вида $P_S = (p_{ij})$ размерности $n \times n$, $i, j = \overline{0, n-1}$, а π_0 − вектор начального распределения.

Автономным вероятностным автоматом будем называть систему [6]

$$(\, S, \varphi(s' / s)), \tag{2}$$

где $s, s' \in S$, $\varphi(s', s)$ − функция переходов, заданная стохастической матрицей P_S .

ЦСМ$_r$ имеет период $r > 1$, а ее состояния подразделяются на r циклических классов − $\overline{S_1}$, $\overline{S_2}$, …, $\overline{S_r}$: ЦМ движется по циклическим классам в определенном порядке, возвращаясь в класс с начальным состоянием через r шагов [1], при этом $n_i = \left|\overline{S_i}\right|$, $i = \overline{1, r}$, $\sum_{i=1}^{r} n_i = n$.

В качестве АММ(P), порождающей ЦМ, будем рассматривать автомат вида (2), заданный на основе (1) по алгоритму разложения ЭСМ P [6]. На АММ(P) накладывается ограничение: значения положительных элементов P кратны заданной величине D.

Если АММ определена как «черный ящик», когда неизвестна функция перехода $\varphi(s', s)$ АММ, но известны последовательности $\hat{S}(N)$ то можно рассматривать задачи идентификации АММ по реализациям ЦМ.

Для каждой последовательности $\hat{S}(N)$ необходимо определить значения α, вероятности того, что $\hat{S}(N)$ сгенерирована на основе АММ(P), где ЭСМ P принадлежит заданному подклассу Q_L.

Для решения поставленной задачи производится модификация модели распознавания, предложенного в [5] в рамках скрытой марковской модели, и адаптация алгоритма «прямого-обратного» хода для определения $P(\hat{S}(N) \mid \text{АММ}(P \in Q_L))$ − вероятности того, что заданная последовательность $\hat{S}(N)$ порождаема автоматной марковской моделью, где ЭСМ P размерности $n \times n$ принадлежит заданному подклассу Q_L и имеет максимальную энтропию.

Алгоритм включает три этапа:

1) инициализация: $\alpha_1(i) = \pi_0(i) \cdot z_i$, $i = \overline{1, m}$, $z_i = \begin{cases} 1: & s(t+1) = s_j \\ 0: & \text{иначе} \end{cases}$;

2) индукция: $\alpha_{t+1}(j) = \left[\sum_{i=1}^{m} \alpha_t(i) \cdot p_{ij}\right] \cdot z_j$, $t = \overline{1, N-1}$, $j = \overline{1, m}$;

3) находим $P(\hat{S}(N) \mid \text{АММ}(P)) = \alpha_N(s(N))$.

Для идентификации последовательности, k состояний которой скрыто от наблюдения − $\hat{S}_k(N)$, при выполнении этапа 2 вычисления й $\alpha_{t+1}(i)$, $t = \overline{1, N-1}$, $i = \overline{1, m}$, имеет место выражение: $\alpha_{t+1}(j) = \left[\sum_{i=1}^{n} \alpha_t(i) \cdot p_{ij}\right] \cdot z'_j$, $z'_j = \begin{cases} 1: & s(t+1) - \text{скрыто} \\ z_j: & \text{иначе} \end{cases}$. Кроме того, если $s(N)$ скрыто от наблюдения, то вероятность $P(\hat{S}_k(N) \mid \text{АММ}(P)) = \sum_{i=1}^{m} \alpha_N(i)$.

Рассмотрим пример применения указанного алгоритма для идентификации заданных последовательностей [7].

Пусть задана ЭСМ P размерности $m = 3$. P имеет максимальную энтропию и определяет АММ(P):

$$P=\begin{pmatrix}0,(3) & 0,(3) & 0,(3)\\ 0 & 0{,}5 & 0{,}5\\ 0{,}5 & 0 & 0{,}5\end{pmatrix},\quad \pi_0(i)=0,(3),\quad i=\overline{1,3}.$$

Первая последовательность определена полностью: $\hat{S}(4)=s_1s_3s_2s_1$.

Согласно предложенному алгоритму, $\alpha_1=(0,(3)\quad 0\quad 0)$, $\alpha_2=(0\quad 0\quad 0,(1))$, $\alpha_3=(0\quad 0\quad 0)$, $\alpha_4=(0\quad 0\quad 0)$. В результате $P(\hat{S}(4)\,|\,\text{АММ}(P))=0$. Данный результат позволяет сделать заключение о том, что первая последовательность не может быть сгенерирована на основе заданной АММ(P).

Вторая последовательность определена полностью: $\hat{S}(4)=s_1s_2s_3s_1$.

Согласно алгоритму, $\alpha_1=(0,(3)\quad 0\quad 0)$, $\alpha_2\approx(0\quad 0,(1)\quad 0)$, $\alpha_3\approx(0\quad 0\quad 55{,}6\cdot 10^{-3})$, $\alpha_4\approx(27{,}8\cdot 10^{-3}\quad 0\quad 0)$. В результате $P(\hat{S}(4)\,|\,\text{АММ}(P))\approx 27{,}8\cdot 10^{-3}$.

Третья последовательность определена не полностью: $\hat{S}_2(4)=s_1(?)s_3(?)$.

Согласно алгоритму, $\alpha_1=(0,(3)\quad 0\quad 0)$, $\alpha_2\approx(0,(1)\quad 0,(1)\quad 0,(1))$, $\alpha_3\approx(0\quad 0\quad 148\cdot 10^{-3})$, $\alpha_4\approx(74{,}1\cdot 10^{-3}\quad 0\quad 74{,}1\cdot 10^{-3})$. В результате $P(\hat{S}_2(4)\,|\,\text{АММ}(P))\approx 148\cdot 10^{-3}$.

Также предложена модификация алгоритма «прямого-обратного» хода для АММ, где $P\in P_n(\text{ЦСМ}_\text{г})$, который, аналогично приведенному выше алгоритму, включает этапы 1) − 3). Этапы повторяются для каждой P $S(n,r)$ раз, где $S(n,r)$ − число Стирлинга II рода.

В результате, модифицированный алгоритм «прямого-обратного хода» применим к решению задач распознавания АММ, постановки которых приведены в [7-10]. Идентификация возможна как для ЦМ указанного класса, все состояния которых наблюдаемы в полном объеме, так и для ЦМ, часть состояний которых скрыты от наблюдения. Представленная модель позволяет количественно оценить вероятность идентификации ЦМ на предмет возможности генерирования заданной АММ.

Список литературы

1. *Романовский, В.И.* Дискретные цепи Маркова. М.: Гостехиздат, 1949. 436 с.

2. *Бухараев Р.Г.* Вероятностные автоматы. Казань: Изд-во КГУ, 1977. 247 с.

3. *Захаров В.М., Н.Н. Нурмеев, Ф.И. Салимов и др.* К задаче дискриминантного анализа автоматных марковских моделей // Вестник КГТУ им. А.Н. Туполева. 2001. № 3. С. 37–39.

4. *Поспелов Д.А.* Вероятностные автоматы. М.: Энергия, 1970. 88 с.

5. *Рабинер Л.Р.* Скрытые марковские модели и их применение в избранных приложениях при распознавании речи // ТИИЭР. 1989. Т. 77. № 2.

6. *Кемени Дж., Снелл Дж.* Конечные цепи Маркова. М.: Наука, 1970. 272 с.

7. *Шалагин С.В., Нурутдинова А.Р.* Модель идентификации конечных простых однородных цепей Маркова // Сб. трудов V науч.-техн. конф. преподавателей, научных работников и аспирантов с междунар. участием «Актуальные проблемы и перспективы развития гражданской авиации», 22–24 марта 2016 г. Иркутск: Иркутский филиал МГТУ ГА, 2016. С. 116–122.

8. *Шалагин С.В.* Идентификация источника последовательности на основе скрытых цепей Маркова // Материалы Всеросс. науч.-практ. конф. с междунар. участием «Новые технологии, материалы и оборудование российской авиакосмической отрасли», 10-12 августа 2016 г. / Сб. докладов. Т. 2. Казань: Изд-во Академии наук РТ, 2016. Т. 2. С. 292–297.

9. *Shalagin S.V., Nurutdinova A.R.* Identification Algorithms of Simple Homogeneous Markov Chains of Cyclic Class and Their Complexity Analysis // International Journal of Pharmacy and Technology. 2016. Vol. 8(3). P. 14926–14935.

10. *Shalagin S.V., Nurutdinova A.R.* Identification of Markovian Automata Sub-classes // International Journal of Pharmacy and Technology. 2016. Vol. 8(3). P. 15327–15337.